The authors would like to sincerely thank Reviewer #2 for their insightful comments and their time for reviewing this manuscript. We have addressed their comments below and believe their feedback has strengthened the science of this manuscript.

Reviewer 2

This paper describes the analysis by continuous flow techniques of two upper sections of ice core taken from the Allan hills region of Antarctica. Ice cores from this region have sampled by far the oldest dateable ice in Antarctica, with ice as old as 6Ma reported. The significance of this is the access to climate and atmospheric signals far older than the stratigraphic ice cores such as from EDC, Dome Fuji and Vostok. Their value then is enabling data from, and insights into, earlier significant climate epochs such as the mid-Pleistocene transition when the glacial cycle changed pacing from ~40ka to ~100ka, and the earlier mid-Pliocene warm period.

The authors, and indeed the wider COLDEX community would recognise that dating these older cores is problematic, and that the proxy signals are difficult to attribute and understand due to the accumulation and complex flow regime of the ice itself, with stagnation, overturning and upwelling of the ice as it closes on to the Transantarctic mountains.

This paper takes some novel approaches to the analysis of climate proxy data from two fairly shallow sections of cores (46 and 69m – the upper, least damaged, sections of deeper ice cores recovered from the Allan Hills region) which nevertheless include ice that may be as old as 1.2 Ma. It examines evidence for biasing of the proxy data towards warm interglacial periods that has previously been reported.

The paper is well written, and interesting, and certainly appropriate for Climate of the Past. It does tend towards descriptive statistics and sometimes simplistic explanations of likely causes of the observations, but there are some interesting approaches to looking at the data that others will likely take up with other cores.

Generally, the paper is fine, though there are a mixture of minor and perhaps more substantial comments below that the authors may wish to address. There is a tendency to write off anything that doesn't fit the general theme of the paper as 'contamination' (e.g. in discussing methane data), or 'stratigraphic inconsistencies' which probably captures the main reason why these cores are difficult to characterise! And when the two cores show different characteristics, then 'differing flow regimes' can be brought to bear.

The major observations of biasing of the ALHIC cores towards interglacials might be more robust if the choice of data sampling from the EDC core was better justified. The EDC data is sampled by taking 1000 random mean data points from one meter depth intervals. The use of data on a depth scale already introduces a bias of EDC data points taken from the glacial period given that the last interglacial spans a significant portion of the (depth) record. A quick histogram of the EDC isotope data in binned 1m intervals, and an eyeball analysis suggest about two thirds of the record is glacial. So, it might not be surprising to find the ALHIC records biasing towards the interglacial given that they have (presumably) a quite different age-depth distribution. It is also worth pointing out that the variance of the EDC data differs across the climate states, which would impact on some of the analysis here.

We thank the reviewer for their evaluation. In response to this comment, section 2.4 has been modified to add this section:

“The purpose of the EDC comparison is to assess whether discontinuous blue ice at Allan Hills reliably preserves environmental signals and glacial-interglacial variability. ALHIC 2201 and 2302 do not have a robust age scale due to discontinuous stratigraphy in conjunction with limited discrete 40Ar dates. Therefore, the EDC comparative analyses are conducted with all EDC data resampled on a 1 m binned depth scale rather than an age scale. Comparison between these two records on a depth scale allows for interpretation of the distribution of environmental variability and the depth binning of the EDC record limits the influence of sampling resolution biases as a result of uneven depth sampling in the EDC datasets. The continuous 800 kyr EDC record is an appropriate benchmark for the ALHIC CFA records in that most of the estimated ages fall within this age range (see Section 3.1) and the distributional comparison of the environmental variables is independent of the continuity of the record. Due to regional differences between Victoria Land and the East Antarctic Plateau, a mean offset is applied to the Allan Hills water isotope and dust values for better comparison to EDC.

We would like to note that not all analyses were randomly sampled for 1000 datapoints, this was just for Figure 4 for the correlation analyses. For all analyses, EDC data were binned at 1 m depth intervals unless otherwise stated.

In addition, we believe acknowledging 2/3 of the EDC record is glacial is part of what motivates our analyses; these features of the EDC dataset are part of the benchmark we want to test our ALHIC distributions against. We want to understand if you drill at the Allan Hills, do we recover an archive that is similar in nature to those of ‘traditional’, continuous cores like EDC or does it have a different behavior? We acknowledge that EDC does have different variability across climate states but because we do not have a proper age scale for the Allan Hills, it is most robust to compare to the full EDC record on a depth scale (rather than age scale).

The CFA (continuous flow analysis) technique applied to the two cores is very well described, with a clear explanation of the analytical procedures and methods of calibration, and good figures illustrating the methodology. In terms of modern ice core CFA techniques, particularly those of the European groups, the range of analytes is limited, with only methane, microparticles (dust) and the stable water isotopes delD and del18O measured. But, given the later scope of the manuscript, which examines whether Allan Hills cores can reliably capture long-term trends in paleoclimate, this is not a limitation, but a sensible selection of primary proxies with which to test whether blue ice cores from the Allan Hills adequately record climate data.

There are a few specific remarks on the analytical section:

L66: the sentence “….and glacial ice to the average are yet to be fully” is incomplete.

Thank you for noticing this error. We added “understood” to the end of this sentence.

L195: in describing the validation of the methane calibration, comparison is made between the CFA methane, and measurements on discrete samples taken from the two cores. This showed a mean offset of -0.29 ppb, with a precision of +/- 53.2 ppb. I can’t help feeling that this almost perfect minimal offset is either luck or judicious rejection of data that do not compare well. This can anyway be seen in the figure 3, where the comparison is poor in the upper 10m of the 2302 core (and rejected from the comparative statistics). I’d like to see a little more written detail on the comparison between discrete and CFA data – actually figure 3 looks pretty good for the 2201 core, so I’m surprised at the poor precision figure rather than any offset.

In re-evaluating the CFA and discrete measurement offsets, we found that the numbers reported need to be updated. The mean and median offset is -6.55 and 2.31 ppb, respectively, with a standard deviation of 57.67 ppb. These large offsets, as stated in the text, are likely due to the depth register issues with ALHIC 2201 due to ice fracturing in transit before processing in conjunction with stratigraphic offsets across the slab of the ice core due to apparent dipping layers. As seen in Figure 3 and 5, many of the discrete methane measurements appear to match the CFA data quite well.

Discarding the upper 10 m is due to surface cracking, a well-known problem at this site and other Blue Ice Areas. We now cite Baggenstos et al., 2017 and Turney et al., 2013 in Section 2.2.2 for more context.

To add more detail for the readers, we added the precision of the discrete methane measurements at OSU are 2.7 ppb (described in Epifanio et al., 2020, *Climate of the Past*) and how the offsets were calculated in Section 2.2.2:

"The offset was calculated by averaging all CFA CH_4 measurements falling within each discrete sample's top and bottom depth interval and subtracting this CFA mean from the corresponding discrete CH_4 measurement."

L206: Note is made that the calibration of the stable isotopes was carried out at the start and end of each day of analysis, but that only an overall campaign calibration was used to calibrate the data. I've no objection to this, but it would have been useful to be given the number of days and the overall period of the campaign. Instrument stability over a few consecutive days of analysis would be more convincing than if the campaign had continued over several weeks. Could you include details of the campaign period in the paragraph L221 et seq?

The following will be added to Section 2.2.2 to provide clarity of the duration of the campaign:

The ALHIC 2201 solubility offset was 0.120 ± 0.016 across eight campaign days between March 4-15th 2024 and the ALHIC 2302 the methane solubility offset was 0.058 ± 0.006 across six campaign days in March 24-26 and April 2, 4, and 15th 2025.

The following has been added to address the water isotope stability over the course of the campaign:

"Calibration standards were run at the beginning and end of each day. The calibrations across all days (separate for each campaign) were combined to calculate a global linear calibration for δ^2H and $\delta^{18}O$, and the slope and intercept were compared to a daily δ^2H and $\delta^{18}O$ calibration using a two-tailed t-test. The daily calibrations were not significantly different from the global calibrations and therefore a global calibration was used."

While commenting on the analytical technique, I'd remark that the figures A1 and A2 are quite convincing. I'd remove the connecting lines across the data gaps in the dust data (e.g. around 42 and 44m) in A1, and for all three species plotted in A2. The dust measured at the two institutions is remarkably similar, given the nature of ice selection and preparation for CFA analysis. Figure A2 is also convincing too; there are small offsets in the methane which I'd expect given the nature of sample preparation (removing contaminated ice from the ends of short CFA sections) and the difficulty of precise depth registration.

Thank you for noticing that the gaps in the data should be removed from the plots. These have been removed in the next iteration of figures.

Figure A3 and the relevant sections of text give a good explanation of the response times for each analyte in the CFA flow path and provide good evidence for the claimed spatial resolution of the data.

No response seems needed here from us.

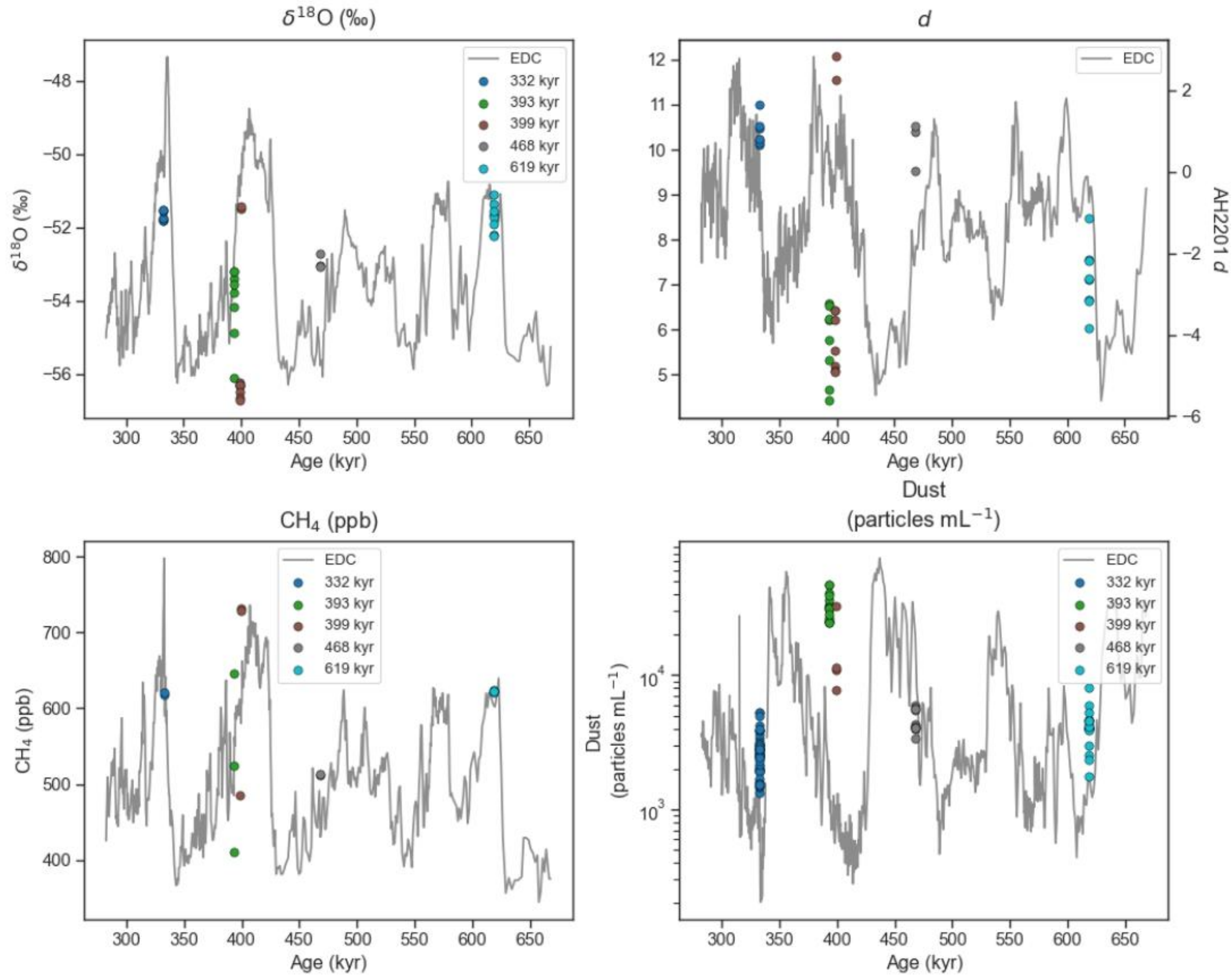
Section 2.3 describes the dating of the ice sections. I'll accept that dating via the 40Ar technique is always going to be difficult and, unlike for example the ~800ka European and Japanese cores, will never have the precision possible where the ice is stratigraphically preserved. But, the ice obtained in the Allan Hills is so much older, so offers climate/gas data that is just never going to be possible from the stratigraphically preserved cores. So, dating the ice is critical to placing the climate data in time. The age uncertainty from the 40Ar method is noted as +/- 11% (L271) or +/- 60ka (L272). However, although the age uncertainty is given, often in the text the ages seem to be taken with more confidence than might be justified.

The clearest evidence that the dating of the ice must be wrong in places is shown in figure A4. There are several places where the methane and del18O indicate the ages of the ALHIC cores must be wrong. Given that methane is globally mixed (as mentioned in the text, L370), there are dated methane values that do not match the methane of the same age in the EDC core (examples at around 150, 350, 390, 550ka in 2201, and 270, 360, 405ka in 2302) with a similar mismatch in the del18O at the same age points. I accept that the dating precision of +/- 60ka could move many of these points laterally in time to match with EDC, but the very fact you have clusters of points of similar age and methane/del18O suggests that the age precision is better than +/-60 ka.

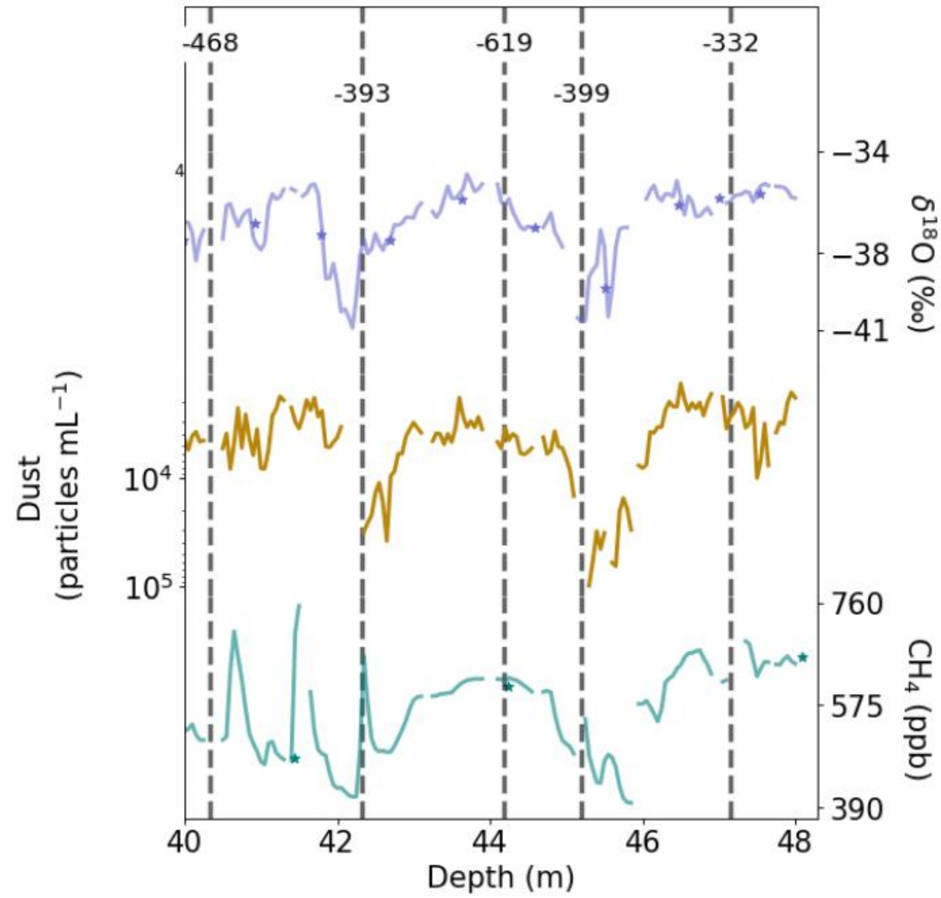
The worry is that errors in ages of the ALHIC cores impact some of the data analysis across the remainder of the manuscript. As an example, figure 5 describes a section of potential folding in core 2201 with an age of 619 ka bracketed by ages 393 and 399 ka – sections that already look incorrectly dated when looking at the methane, d and del18O data around 390 ka in figure A4.

We agree that there are likely dates that do not reflect the true age (and therefore do not match the EDC record) and this is discussed in lines 320-330 as it could be due to layer mixing, for example.

When taking a closer look at only the ages in the proposed fold on Figure A4, we see that the ages before and in the middle, but not on the peaks, correspond with the EDC record well. The datapoints on the 'peaks' seen on the potential fold correspond to ages 393 and 399 kyr. These ages do match somewhat well for methane and $\delta^{18}O$ but not *d* or dust. This may be indicative that this ice is likely both mixed and potentially from a different origin than the surrounding ice packets. A figure here is included for clarity on the section from 40-48 m depth in ALHIC 2201, but will not be included in the manuscript. We expand our explanation of the proposed fold to encompass this explanation. An additional figure added here is a modification of Figure 5B that shows the geochemical measurements with their associated ages for the potential fold that include more ages (and depths).



δ^{18}O (‰)
EDC
332 kyr
393 kyr
399 kyr
468 kyr
619 kyr
δ^{18}O (‰)
Age (kyr)
d
AH2201 d
CH$_4$ (ppb)
Dust
(particles mL^{-1})



We changed the second paragraph of Section 3.1 to the following:

The observed multi-measurement geochemical data in conjunction with ^{40}Ar dating, provide evidence of a potential meter-scale fold in ALHIC 2201 between ~42 and 45 m, where two similar ages occur at ~42 and 45 m (393 and 399 kyr, respectively), surrounding a dissimilar age at 44 m (619 kyr) (Fig. 5B). However, these ages have uncertainties of ~60 kyr, and individual measurements may reflect mixing of ice of different ages rather than the true age of the sampled ice. Figure 5 provides additional context for evaluating these ages by comparing the ALHIC 2201 geochemical measurements with the EDC record. The ages associated with the similar-age peaks (393 and 399 kyr) do not correspond well with the EDC data for dust or *d* (but align relatively well for CH4 and δ^{18}O), whereas the surrounding ice packets (468, 619, and 332 kyr) show stronger alignment. Nevertheless, the covariation and mirrored sequence observed among all three geochemical measurements are consistent with a potential fold in this interval. Kirkpatrick et al. (2025) finds that layer dip angles in the upper layers of ALHIC2201 are consistent with depth. This result is supported by visible, near-parallel englacial layers in radar data near this site (Nesbitt and Brook, 2023). Kirkpatrick et al. (2025) suggested that the apparent contradiction between parallel layering and stratigraphically disturbed age scales may be explained by recumbent folds. In this scenario, folding would occur in place or upstream and be drawn into parallel layers by ice flow and thinning.

L 274: 15cm of adjacent ice – can you clarify this sentence? I assume it is horizontally adjacent, rather than adjacent in the vertical sense. Given that elsewhere in the manuscript (Lines 409 et

seq.) the dipping of the ice layers is given some significance, is it not fair to say that even a horizontally adjacent piece of ice might have a different age? But, then you could point to Figure A2 to show how reproducible are the analyses taken from presumably two adjacent CFA sections of the same core. Though this might then undermine your argument about signal smoothing due to the dipping layers described in L417 et seq.

It is true that the dipping layers mean that analyses from one section of ice may not represent exactly the same layers as analyses taken at the same measured depth from a different side of the core. This is a recurring challenge in Allan Hills ice cores, where co-registered depth alignment is a highly intricate process requiring careful ice cutting and preparation and alignment with 3-D ECM data where available. Additionally, the dipping angle is not homogeneous throughout the core, further complicating layer alignment. In Figure A2, the dipping layers are approximately 20–30°, which could result in offsets of approximately 5–10 cm between horizontally adjacent sections. Because the ^{40}Ar samples cover a depth of 15-20 cm and ~15 cm across the core (from the other side of the BID core), the offset across the core from the CFA to the ^{40}Ar piece would be on the order of only a 3-5 cm offset.

We will revise L274 as follows:

"The samples used for dating were from an adjacent (i.e., approximately the same depth interval) piece of ice from the BID core using an 800 g sample."

L357: you state all geochemical measurements except methane in 2201 have lower variance than EDC. Yet table 1 shows higher variance for 2201 dust and d than EDC.

Thank you for noticing these details. We will note that ALHIC 2201 *d* also has higher variance than EDC. The higher variance for ALHIC 2201 dust was incorrect in the table and has been updated. The number should read 1.0×10^8

L426: rather than "over three times narrower than that of" you could say with more clarity "is less than a third of".

Thank you for this suggestion. We have implemented this change.

L458: states ice cores only record environmental conditions only during snowfall. This is not correct: dry deposition of many species is likely, especially in low accumulation areas. Specifically, dust can be deposited independently of accumulation (as the dust on the hood of my car during a dry spell attests). Of the proxies described in this paper, methane is trapped independent of snowfall – it is the depth of pore close off and trapping rather than the magnitude of the methane signal that changes with accumulation rate.

Thank you for this important note. We will revise line 458 as follows:

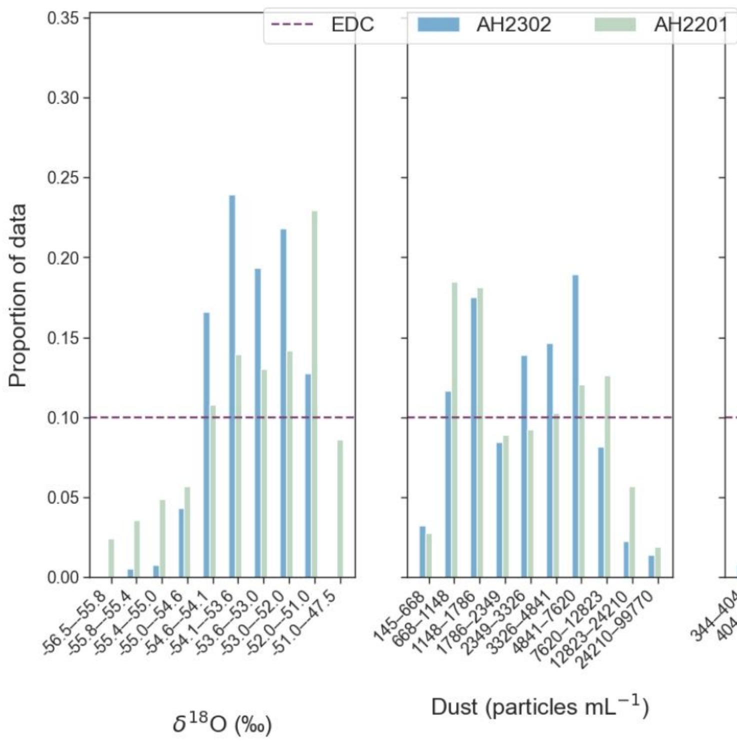
"Ice cores inherently overrepresent periods of higher snow accumulation, such as interglacials, because more ice is added and preserved than periods of low accumulation."

Section 3.3.1 takes a decile-based approach to testing the biasing in the climate capture in the ALHIC cores. It's quite a simple, but clever approach and figure 7 does seem to justify the observation in the text that, for example, the majority of the ALHIC methane data fall above the median EDC methane value, biasing ALHIC methane towards the interglacial climate

states. However, as previously noted, this might arise from the fact that EDC was sampled on 1m depth intervals, and the glacial periods are over-emphasised in EDC data on a depth scale.

The findings in this analysis are similar when EDC is age-binned instead of depth-binned, providing evidence this method is robust to the sampling method. A figure of the EDC data binned by 1000 kyr.

Note that Allan Hills at 5 cm resolution, like we use, is also effectively binned on a depth scale. There is no Allan Hills age scale available, so this is by necessity.



Section 3.3.2 and Appendix B is a quite clever analysis of the data distribution and smoothing. It is more robust in statistical technique than some of the more descriptive statistics appearing in the main body of the paper. While I could not follow all the detail (what is a Wasserstein difference – this could be explained), it was an interesting (new?) way of using a remote dataset (the EDC core) with resampling to examine biasing and smoothing in the ALHIC cores. It may identify the differences between ALHIC and EDC cores, but surely does not explain the climatic/glaciological origin of the differences. And how could it? The age models of ALHIC cores are not well constrained, and flow regimes are so strikingly different. Nevertheless, a very interesting approach.

Thank you for highlighting the importance of this analysis and its potential interest to readers. We agree that this analysis offers important context to the observed patterns as seen in the descriptive analyses. We now include a brief definition in the main text of the Wasserstein distance in the first paragraph of Section 3.3.2.

However, we would like to retain the current format of the manuscript, with the main focus being on the descriptive analyses and maintaining the discussion in Section 3.2.2 and leaving most of the methods and a supplemental figure in Appendix B. We believe that these descriptive analyses of the distributions and variances are a fundamental observation needed for understanding the Allan Hills archive. The aim of the additional quantitative analyses is to offer an exploration into understanding potential drivers of the distributional patterns observed in the archives. Because this is an exploratory analysis and not a fundamental character of the archive itself, we would like to maintain the structure of the manuscript to allow these analyses to be the focus of the main conclusion of the manuscript.

To reduce the number of descriptive analyses we will move Figure 4 (the correlation figure) to Appendix A, as it does not have a substantial implication for the manuscript.